

COMPUTABLE ABSOLUTELY PISOT NORMAL NUMBERS

MANFRED G. MADRITSCH, ADRIAN-MARIA SCHEERER, AND ROBERT F. TICHY

ABSTRACT. We analyze the convergence order of an algorithm producing the digits of an absolutely normal number. Furthermore, we introduce a stronger concept of absolute normality by allowing Pisot numbers as bases, which leads to expansions with non-integer bases.

1. INTRODUCTION

In this paper we are interested in simultaneous normality to several bases. In particular, we analyze the order of convergence to normality of an absolutely normal number generated by an algorithm of Becher, Heiber and Slaman (Section 2) and are concerned with normality to real bases. We give an algorithmic construction of a real number that is normal to each base from a given sequence of Pisot numbers (Section 3 and Section 4).

1.1. Normality to a single base. A real number $x \in [0, 1)$ is called *simply normal to base b* , $b \geq 2$ an integer, if in its b -ary expansion

$$x = \sum_{n \geq 1} a_n b^{-n}, \quad a_n \in \{0, \dots, b-1\}$$

every digit $d \in \{0, 1, \dots, b-1\}$ appears with the expected frequency, i.e. the limit

$$\lim_{n \rightarrow \infty} \frac{1}{N} |\{1 \leq n \leq N : a_n = d\}|$$

exists and is equal to $\frac{1}{b}$. The number x is called *normal to base b* , $b \geq 2$ an integer, if in its b -ary expansion all finite combinations of digits appear with the expected frequency, i.e. if for all $k \geq 1$ and all $d \in \{0, \dots, b-1\}^k$,

$$(1.1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} |\{1 \leq n \leq N : (a_n, \dots, a_{n+k-1}) = d\}| = \frac{1}{b^k}.$$

Normal numbers were introduced by Borel [11] in 1909. He showed that almost all real numbers (with respect to Lebesgue measure) are simply normal to all bases $b \geq 2$, thus *absolutely normal* (see Section 1.3). It is a long standing open problem to show that important real numbers such as $\sqrt{2}$, $\ln 2$, e , π , \dots are normal, for instance in decimal expansion. There has only been little progress in this direction in the last decades, see e.g. [2].

Date: October 21, 2016.

2010 *Mathematics Subject Classification.* 11K16 (primary), 11Y16 (secondary).

However, specifically constructed examples of normal numbers are known. Champernowne in 1935 [15] has shown that the real number constructed by concatenating the expansions in base 10 of the positive integers, i.e.

$$0,1234567891011\dots,$$

is normal to base 10. This construction has been extended in various directions (*cf.* Erdős and Davenport [16], Schiffer [30], Nakai and Shiokawa [24], Madritsch, Thuswaldner and Tichy [23], Scheerer [29]).

1.2. Discrepancy of normal numbers. The *discrepancy* of a sequence $(x_n)_{n \geq 1}$ of real numbers is defined as

$$D_N(x_n) = \sup_J \left| \frac{1}{N} |\{1 \leq n \leq N : x_n \bmod 1 \in J\}| - \lambda(J) \right|,$$

where the supremum is extended over subintervals $J \subseteq [0, 1)$ and where λ denotes the Lebesgue measure. A sequence is *uniformly distributed modulo 1* if its discrepancy tends to zero as $N \rightarrow \infty$.

It is known [37] that x is normal to base b if and only if the sequence $(b^n x)_{n \geq 1}$ is uniformly distributed modulo 1. Hence x is normal to base b if and only if $D_N(b^n x) \rightarrow 0$ as $N \rightarrow \infty$. It is thus a natural quantitative measure for the normality of x to base b to consider the discrepancy of the sequence $(b^n x)_{n \geq 1}$.

Answering a question of Erdős, in 1975 Philipp [26] has shown a law of the iterated logarithm for discrepancies of lacunary sequences which implies $D_N(b^n x) = O(\sqrt{\log \log N / N})$ almost everywhere. Recently, Fukuyama [18] was able to determine

$$\limsup_{N \rightarrow \infty} \frac{D_N(b^n x) \sqrt{N}}{\sqrt{\log \log N}} = c(b) \quad \text{a.e.,}$$

for some explicit positive constant $c(b)$. Schmidt [33] showed that there is an absolute constant $c > 0$ such that for any sequence $(x_n)_{n \geq 1}$ of real numbers $D_N(x_n) \geq c \frac{\log N}{N}$ holds for infinitely many N . Schiffer [30] showed that the discrepancies of constructions of normal numbers in the spirit of Champernowne satisfy upper bounds of order $O(\frac{1}{\log N})$. It is an open question whether there exist a real number x and an integer $b \geq 2$ such that $D_N(b^n x) = O(\frac{\log N}{N})$.

1.3. Absolute normality and order of convergence. A number x is called *absolutely normal* if it is normal to any integer base $b \geq 2$. Since normality to base b is equivalent to simple normality to all bases b^n , $n \geq 1$, absolute normality is equivalent to simple normality to all bases $b \geq 2$.

Since most constructions of numbers normal to a single base b are concatenations of the b -ary expansions of $f(n)$, $n \geq 1$, where f is a positive-integer-valued increasing function, they essentially depend on the choice of the base b . Therefore they cannot be used for producing absolutely normal numbers.

All known examples of absolutely normal numbers have been established in the form of algorithms¹ that output the digits of this number to some base one after the other. The first such construction is due to Sierpinski [35] from 1917. This construction was made computable by Becher and Figueira [4] who gave a recursive formulation of Sierpinski's construction. Other algorithms for constructing absolutely normal numbers are due to Turing [36] (see also Becher, Figueira and Picchi [5]), Schmidt [32] (see also Scheerer [28]) and Levin [20] (see also Alvarez and Becher [1]).

There seems to be a trade-off between the complexity of the algorithms and the speed of convergence of the corresponding discrepancies. The discrepancies satisfy upper bounds of the order $O(N^{-1/6})$ (Sierpinski), $O(N^{-1/16})$ (Turing), $O((\log N)^{-1})$ (Schmidt) and $O(N^{-1/2}(\log N)^3)$ (Levin). All algorithms, except the one due to Schmidt, need double exponential many mathematical operations to output the first N digits of the produced absolutely normal number. Schmidt's algorithm requires exponentially many mathematical operations.

No construction of an absolutely normal number x is known such that the discrepancy $D_N(b^n x)$ for some $b \geq 2$ decays faster than what one would expect for almost all x .

In Section 2 we are interested in another construction of an absolutely normal number which is due to Becher, Heiber and Slaman [6]. They established an algorithm which computes the digits of an absolutely normal number in polynomial time. We show (Theorem 2.8) that the corresponding discrepancy is slightly worse than $O(\frac{1}{\log N})$, and that at a small loss of computational speed the discrepancy can in fact be $O(\frac{1}{\log N})$.

1.4. Normality to non-integer bases. Section 3 of the present article treats normality in a context where the underlying base is not necessarily integer. Let $\beta > 1$ be a real number. Expansions of real numbers to base β , so-called β -expansions, were introduced and studied by Rényi [27] and Parry [25] and later by many authors from an arithmetic and ergodic-theoretic point of view.

In the theory of β -expansions it is natural to consider *Pisot numbers* β , i.e. real algebraic integers $\beta > 1$, such that all its conjugates lie inside the (open) unit disc. A real number x is called *normal to base β* , or β -normal, if the sequence $(\beta^n x)_{n \geq 1}$ is uniformly distributed modulo 1 with respect to the unique entropy maximizing measure for the underlying transformation $x \mapsto \beta x \bmod 1$ (see Section 3.1). A real number is called *absolutely Pisot normal* if it is normal to all bases that are Pisot numbers. Since there are only countably many Pisot numbers, the Birkhoff ergodic theorem implies that almost all real numbers are in fact absolutely Pisot normal.

The main result of Section 3 is an algorithm that computes an absolutely Pisot normal number. More generally, for a sequence $(\beta_j)_{j \geq 1}$ of Pisot numbers, we construct a real number x that is normal to each of the bases β_j , $j \geq 1$ (Section 3.3 and Theorem 3.6).

¹With the exception of Chaitin's constant, which is absolutely normal but not computable [14].

Bearing in mind that the set of computable real numbers is countable, we thus show that there is in fact a computable real number that is β_j -normal for each $j \geq 1$.

Our algorithm constructs in each step a sequence of finitely many nested intervals, corresponding to the first finitely many bases considered. This is also the essential idea of the construction of an absolutely normal number by Becher, Heiber and Slaman [6]. We need to establish lower and upper bounds for the length of β -adic subintervals in a given interval to control the number of specified digits when changing the base. However, the equivalence (absolute normality) \Leftrightarrow (simple normality to all bases) does not hold for non-integer expansions. Instead, we argue with the concept of (ε, k) -normality as introduced by Besicovitch [9] and studied in the case of Pisot numbers by Bertrand-Mathis and Volkmann [8].

Our algorithm should be compared to the one due to Levin [20]. While his construction is not restricted to Pisot numbers, it uses exponential sums and is as such not realizable only with elementary operations. The algorithm we present in Section 3 is completely elementary.

In Section 4 we give explicit estimates of all constants that appear in our algorithm. We use a theorem on large deviations for a sum of dependent random variables to give an estimate for the measure of the set of non- (ε, k) -normal numbers of length n (Proposition 4.3). Our approach gives all implied constants explicitly, and as such makes a consequence of the ineffective Shannon-McMillan-Breimann theorem effective. The results of this section might be of independent interest.

1.5. Notation. For a real number x , we denote by $\lfloor x \rfloor$ the largest integer not exceeding x . The fractional part of x is denoted as $\{x\}$, hence $x = \lfloor x \rfloor + \{x\}$. We put $\lceil x \rceil = -\lfloor -x \rfloor$. Two functions f and g are $f = O(g)$ or equivalently $f \ll g$ if there is a x_0 and a positive constant C such that $f(x) \leq Cg(x)$ for all $x \geq x_0$. We mean $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$ when we say $f \sim g$ and $g \neq 0$.

When we speak of *words*, we mean finite or infinite sequences of symbols (called letters) of a certain (specified) set, the alphabet. *Blocks* are finite words. The concatenation of two blocks $u = u_1 \dots u_k$ and $v_1 \dots v_l$ is the block $u_1 \dots u_k v_1 \dots v_l$ and is denoted by uv or $u * v$. If u_i for $i \leq m$ are blocks, $*_{i \leq m} u_i$ is their concatenation in increasing order of i . The *length* of the block $u = u_1 \dots u_k$ is denoted by $\|u\|$ and is in this case equal to k .

We denote by λ the Lebesgue measure.

For a finite set, $|\cdot|$ means its number of elements.

Mathematical operations include addition, subtraction, multiplication, division, comparison, exponentiation and logarithm. *Elementary operations* take a fixed amount of time. The cost of mathematical operations depends on the digits of the input or on the desired precision of the output. Addition or subtraction of two n -digit numbers takes $O(n)$ elementary operations, multiplication or division of two n -digit numbers takes $O(n^2)$ elementary operations, and to compute the first n digits of \exp and \log takes $O(n^{5/2})$ elementary operations. These estimates are crude but sufficient for our purposes.

The complexity of a computable function f is the time it takes to compute the first N values $f(i)$, $1 \leq i \leq N$. The algorithm we analyze outputs the digits of a real number X to some base. By the complexity of the algorithm we mean the time it takes to output the first N digits of X to some base.

2. DISCREPANCY

In this section, we analyze the speed of convergence to normality of the absolutely normal number produced by the algorithm by Becher, Heiber and Slaman in [6]. We follow the notation and terminology therein.

2.1. The Algorithm.

Notation. A t -sequence is a nested sequence of intervals $\mathbf{I} = (I_2, \dots, I_t)$, such that I_2 is dyadic and for each base $2 \leq b \leq t-1$, I_{b+1} is a $(b+1)$ -adic subinterval of I_b such that $\lambda(I_{b+1}) \geq \lambda(I_b)/2(b+1)$.

Let $x_b(\mathbf{I})$ be the block in base b such that $0.x_b(\mathbf{I})$ is the representation of the left endpoint of I_b in base b . In each step i , the algorithm computes a sequence $\mathbf{I}_i = (I_{i,2}, \dots, I_{i,t_i})$ of nested intervals $I_{i,2} \supset \dots \supset I_{i,t_i}$. If $b \leq t_i$, let $x_b(\mathbf{I}_i) = x_{i,b}$ be the base b representation of the left endpoint of $I_{i,b}$ and let $u_{i+1,b} = u_b(\mathbf{I}_{i+1})$ be such that $x_{i+1,b} = x_{i,b} * u_{i+1,b}$.

If u is a block of digits to base b , the *simple discrepancy* of u in base b is defined as $D(u, b) = \max_{0 \leq d < b} |N_d(u)/\|u\| - 1/b|$ where $N_d(u)$ is the number of times the digit d appears in the block u .

Let $k(\varepsilon, \delta, t)$ be the function

$$k(\varepsilon, \delta, t) = \max(\lceil 6/\varepsilon \rceil, \lceil -\log(\delta/(2t))6/\varepsilon^2 \rceil) + 1.$$

From Lemma 4.1 and 4.2 of [6] we further have a function h that counts the number of mathematical operations needed to carry out one step of the algorithm. See also Lemma 2.5.

Input. A computable non-decreasing unbounded function $f : \mathbb{N} \rightarrow \mathbb{R}$ such that $f(1)$ is known and satisfies $f(1) > h(2, 1)$.

First step. Set $t_1 = 2$, $\varepsilon_1 = \frac{1}{2}$, $k_1 = 1$ and $\mathbf{I}_1 = (I_{1,2})$ with $I_{1,2} = [0, 1)$.

Step $i+1$ for $i \geq 1$. Given are from step i of the algorithm values $t_i = v$, $\varepsilon_i = \frac{1}{v}$ and a t_i -sequence \mathbf{I}_i .

We want to assign values to $t_{i+1}, \varepsilon_{i+1}$. If $i+1$ is a power of 2, then we carry out the following procedure.

- We spend i computational steps on computing the first m values of f , $1 \leq m \leq i$.
- We put $\delta = (8t_i 2^{t_i+v+1} t_i! (v+1)!)^{-1}$.
- We try to compute $k(\frac{1}{v+1}, \delta, v+1)$ and $h(v+1, \frac{1}{v+1})$ in i steps each. If we succeed in computing these values, and if additionally

$$(2.1) \quad h(v+1, \frac{1}{v+1}) < f(m)$$

and for each $b \leq t_i$

$$(2.2) \quad \frac{\lceil \log_2(v+1) \rceil k(1/(v+1), \delta, v+1) + \lceil -\log_2(\delta) \rceil}{\|x_{i,b}\|} < \frac{1}{v+1},$$

then we define $t_{i+1} = v+1$ and $\varepsilon_{i+1} = \frac{1}{v+1}$. Otherwise, we let $t_{i+1} = t_i = v$, $\varepsilon_{i+1} = \varepsilon_i = \frac{1}{v}$.

If $i+1$ is no power of 2, then define $t_{i+1} = t_i = v$, $\varepsilon_{i+1} = \varepsilon_i = \frac{1}{v}$.

Furthermore, we compute $\delta_{i+1} = (8t_i 2^{t_i+t_{i+1}} t_i! t_{i+1}!)^{-1}$ and

$$k_{i+1} = \max(\lceil 6/\varepsilon_{i+1} \rceil, \lceil -\log(\delta_{i+1}/(2t_i)) 6/\varepsilon_{i+1}^2 \rceil) + 1.$$

Then we find a t_{i+1} -sequence \mathbf{I}_{i+1} by means of the following steps.

- We let L be a dyadic subinterval of I_{i,t_i} such that $\lambda(L) \geq \lambda(I_{i,t_i})/4$.
- For each dyadic subinterval J_2 of L of measure $2^{-\lceil \log_2 t_i \rceil k_{i+1}} \lambda(L)$, we find $\mathbf{J} = (J_2, J_3, \dots, J_{t_{i+1}})$, a t_{i+1} -sequence starting with J_2 .
- Finally we choose \mathbf{I}_{i+1} to be the leftmost of the t_{i+1} sequences \mathbf{J} considered above such that for each $b \leq t_i$, $D(u_b(\mathbf{J}), b) \leq \varepsilon_{i+1}$.

Output. Let X be the unique real number in the intersection of the intervals of the sequences \mathbf{I}_i . In base b we have $X = \lim_{i \rightarrow \infty} 0.x_{i,b} = 0.*_{i \geq 1} u_{i,b}$. It is the content of Theorem 3.9 in [6] that X is absolutely normal.

2.2. Speed of convergence to normality. In this section we estimate the discrepancy $D_N(b^n X)$ for integer $b \geq 2$. Two factors play a role: How many digits in each step are computed, and how rapidly ε_i decays to zero. By virtue of the algorithm, at least one digit is added in each step, and ε_i can decay at most as fast as $O(\frac{1}{\log i})$. As can be expected from the algorithm, the discrepancy depends both on growth and complexity of f .

It was shown in [6] that to output the first N digits of X , the algorithm requires time $O(N^2 f(N))$.

We begin our analysis by first showing that in each step of the algorithm not too many digits are attached.

Lemma 2.1 (Lemma 3.3 in [6]). *For an interval I and a base b , there is a b -adic subinterval I_b such that $\lambda(I_b) \geq \lambda(I)/(2b)$.*

Lemma 2.2. *If i is large enough, then $1 \leq \|u_{i,b}\| \ll (\log i)^A$ for $A > 3$. Thus $i \ll \|x_{i,b}\| \ll i(\log i)^A$.*

Proof. We assume the base b to be fixed and i large enough such that $t_{i+1} \geq b$. In step $i+1$ we have the following sequence of nested subintervals:

$$(2.3) \quad I_{i,b} \supset \dots \supset I_{i,t_i} \supset L \supset I_{i+1,2} \supset \dots \supset I_{i+1,b}.$$

By Lemma 2.1, and the choice of $I_{i+1,2}$, we know the following lower bounds on the measures of the intervals in (2.3). We have $\lambda(I_{i,t_i}) \geq \lambda(I_{i,b})/(2^{t_i-b} t_i! / b!)$, $\lambda(L) \geq \lambda(I_{i,t_i})/4$, $\lambda(I_{i+1,2}) = 2^{-\lceil \log_2 t_i \rceil k_{i+1}} \lambda(L)$ and $\lambda(I_{i+1,b}) \geq \lambda(I_{i+1,2})/(2^b b!)$. Combining inequalities yields $\lambda(I_{i+1,b}) \geq$

$\lambda(I_{i,b})/(2^{2+t_i}2^{\lceil \log_2 t_i \rceil k_{i+1} t_i!})$. Hence in stage $i+1$ we are adding at most $O(t_i + (\log t_i)k_{i+1} + \log t_i!)$ many digits in base b . The way the algorithm is designed only allows for $t_i = O(\log i)$. The growth of k_{i+1} can be analyzed and is $O(t_i^3 \log t_i)$. Hence in stage $i+1$ at most $O(t_i + (\log t_i)k_{i+1} + \log t_i!) = O((\log i)^A)$ digits are added to the b -ary expansion of X , where $A > 3$ to accommodate all double-log factors.

The lower bound on the number of digits added comes from the fact that by the choice of $I_{i+1,2}$, $I_{i+1,b}$ is strictly smaller than $I_{i,b}$, so at least one digit is added in each stage. \square

Next, we investigate the conditions involving k and h that are responsible for how fast $t_i \rightarrow \infty$ and $\varepsilon_i \rightarrow 0$ with step i of the algorithm. We start by showing that condition (2.2) on k always holds, provided i is large enough. This involves estimating the growth as well as the complexity of k .

Recall that $k(\varepsilon, \delta, t) = \max(\lceil 6/\varepsilon \rceil, \lceil -\log(\delta/(2t))6/\varepsilon^2 \rceil) + 1$.

Lemma 2.3. *Let $v \geq 2$ be an integer and $\delta = (8v2^{2v+1}v!(v+1)!)^{-1}$. Then the growth of $k(\frac{1}{v+1}, \delta, v+1)$ is $O(v^3 \log v)$. Furthermore, $k(\frac{1}{v+1}, \delta, v+1)$ can be computed in $O(v^2(\log v)^2)$ elementary operations.*

Proof. We have for the growth

$$\begin{aligned} k\left(\frac{1}{v+1}, \delta, v+1\right) &= \max(\lceil 6(v+1) \rceil, \lceil \log(2(v+1)8v2^{2v+1}v!(v+1)!)6(v+1)^2 \rceil) + 1 \\ &\leq 6(v+1)^2 (\log(16v(v+1)) + (2v+1) \log 2 + \log v! + \log(v+1)!) + 2 \\ &= O(v^2(\log v + v + v \log v)) \\ &= O(v^3 \log v). \end{aligned}$$

Since in the expression for k we are rounding, the most relevant part is the computation of the significant digits of $\log(16v(v+1)2^{2v+1}v!(v+1)!)$. The argument of this expression is computable with $O(v^2(\log v)^2)$ elementary operations and has $O(v \log v)$ many digits. We only need to compute $O(\log v)$ many digits of the logarithm, which takes another $O((\log v)^{5/2})$ elementary operations. In total this are $O(v^2(\log v)^2)$ many elementary operations. \square

Corollary 2.4. *For i to be large enough, condition (2.2) on k is always satisfied, i.e. for each $b \leq t_i$*

$$\frac{\lceil \log(v+1) \rceil k(1/(v+1), \delta, v+1) + \lceil -\log(\delta) \rceil}{\|x_{i,b}\|} < \frac{1}{v+1}$$

where v is such that $t_i = v = 1/\varepsilon_i$.

Proof. This is a consequence of $k(1/(v+1), \delta, v+1) = O(v^3 \log v)$, $\log(1/\delta) = O(v \log v)$, $\|x_{i,b}\| \gg i$ and $v = t_i = O(\log i)$ by the way the algorithm is designed. \square

Now we investigate condition (2.1) on h involving f . The function h counts the number of mathematical operations needed to carry out one step of the algorithm. We want to know an upper bound for the growth of h .

Lemma 2.5. *With $t_i = \frac{1}{\varepsilon_i} = O(\log i)$ we have*

$$h(t_i, \varepsilon_i) = O(i^{\log^4 i}).$$

This upper bound for h can be computed with i elementary operations, provided i is large enough.

Proof. The function h decomposes as $h = h_*(h_1g + h_2 + h_3 + h_4)h_0$ as can be seen from the proof of Lemma 4.2 in [6]. Here:

- g (from Lemma 4.1 in [6]), is the minimum number of digits sufficient to represent all the endpoints of the intervals that we are working with in one step (squared). We know from Lemma 2.2 that $g = O(i^2(\log i)^{2A})$ for $A > 3$.
- It takes h_1g many mathematical operations to find a t_{i+1} -sequence for each J_2 . We have $h_1 = t_{i+1}$.
- h_2 is the number of mathematical operations needed to compute the base b representation $u_b(\mathbf{J})$ for each $2 \leq b \leq t_i$. We have $h_2 \leq \lceil \log_2 t_i \rceil k_{i+1}$.
- h_3 counts the number of mathematical operations needed to compute thresholds of the form $(1/b + \varepsilon_{i+1})\|u_b(\mathbf{J})\|$. We have $h_3 = t_i$.
- h_4 comes from counting occurrences of digits in $u_b(\mathbf{J})$ and comparing with the previously computed thresholds. We have $h_4 \ll t_i(\lceil \log_2 t_i \rceil k_{i+1})^2$.
- h_* is the maximum number of iterations it takes to find a suitable t_{i+1} -sequence. There are $2^{\lceil \log_2 t_i \rceil k_{i+1}}$ many different subintervals J_2 of L , hence $h_* = 2^{\lceil \log_2 t_i \rceil k_{i+1}}$. With $k_{i+1} = O(\log^4 i)$ we obtain $h_* = O(i^{\log^4 i})$.
- Finally, the function h_0 is the number of elementary operations needed to carry out each mathematical operation in one step of the algorithm. Since all values that appear in the calculations of one step of the algorithm are at most exponential in t_i which is at most of order $\log i$, and because the number of elementary operations involved depends only on the number of digits of the numbers involved, h_0 is at most of order $\text{poly}(\log i)$.

These bounds can be seen from Lemma 4.1 and Lemma 4.2 in [6]. Combining them gives $h = O(i^{\log^4 i})$.

Remark that, when t_i is bounded by a slower growing function in i such as $\log \log i$, then the significant term in h comes from g and is a power of i . Otherwise h_* is the significant term.

For the complexity of the upper bound for h , note that $i^{\log^4 i}$ can be computed in a power of $\log i$ many elementary operations, so certainly with i elementary operations when i is large enough. \square

Lemma 2.5 has the following two immediate corollaries for the speed of convergence to normality of Becher, Heiber and Slaman's algorithm.

Proposition 2.6. *Becher, Heiber, Slaman's algorithm achieves discrepancy of $D_N(b^n X) = O(\frac{1}{\log N})$ for f computable in real-time with growth $f \gg i^{\log^4 i}$. In this case, the complexity is $O(i^{2+\log^4 i})$.*

Proposition 2.7. *If f is a polynomial in i of degree d , then the complexity of X is $O(N^{d+2})$ but the discrepancy of $(b^n X)_{n \geq 0}$ is $D_N(b^n X) = O_d(\frac{1}{(\log N)^{1/5}})$.*

Proof. These corollaries follow by observing that the complexity of f is such that f is for large enough i computed up to the actual value $f(i)$ (i.e. $m = i$) and that either the condition on h , (2.1), is satisfied, hence the discrepancy is optimal, or that condition (2.1) is only satisfied for $e^{(\log i)^{1/5}}$ of the values that it is checked for. \square

In a similar manner, using Lemma 2.5, one can show quantitatively how growth and complexity of f influence the discrepancy (and the complexity) of Becher, Heiber, Slaman's algorithm. This can be done for example by measuring complexity and growth of f in the following (crude) way. We denote by $\log_{(k)}$ and $\exp_{(k)}$ the k times iterated logarithm or exponential where $\exp_{(k)} = \log_{(-k)}$, and $\exp_{(0)} = \log_{(0)} = id$. Let c be the integer such that in i elementary operations f can be computed up to a value $f(m)$ with $m \sim \log_{(c)} i$. Let g be the integer such that f grows as $f \sim \exp_{(g)} i$. We allow $g \in \mathbb{Z}$ but c is non-negative.

Theorem 2.8. *Assume f is such that the integers c and g above can be defined. Then Becher, Heiber, Slaman's algorithm computes an absolutely normal number X such that for any base $b \geq 2$,*

$$(2.4) \quad D_N(b^n X) = O\left(\frac{1}{(\log_{(1-g+c)} N)^{1/5}}\right)$$

if $1 - g + c > 0$, and

$$(2.5) \quad D_N(b^n X) = O\left(\frac{1}{\log N}\right)$$

otherwise.

Proof. We have $h \ll \max(\text{poly}(i), e^{t_i^5})$ and $t_i \ll \log i$ by the way the algorithm is defined. t_i only increases if i is a power of two and if $h \leq f(m)$. The latter condition is satisfied for all i large enough if $g - c \geq 1$, and for all i (that are powers of two) that satisfy $i \ll \exp((\exp_{g-c-1}(i))^{1/5})$. With $1/t_i = \varepsilon_i$ this gives in this case an upper bound for the discrepancy of order $1/(\log_{(1-g+c)} N)^{1/5}$. \square

3. ABSOLUTELY PISOT NORMAL NUMBERS

In this section, we give an algorithmic construction of a real number that is normal to each base from a given sequence of Pisot numbers. For more information about β -expansions and β -normal numbers see for example the book [13]. We have partly followed the notation in [8].

3.1. β -expansions of real numbers. Let $\beta > 1$ be a real number. Then each real number $x \in [0, 1)$ has a representation of the form

$$(3.1) \quad x = \sum_{i=1}^{\infty} \varepsilon_i \beta^{-i},$$

with integer digits $0 \leq \varepsilon_i < \beta$. One way to obtain such a representation is the following. Let T_β be the β -transformation $T_\beta : [0, 1) \rightarrow [0, 1)$, $x \mapsto \beta x \pmod{1}$. Then $\varepsilon_i = \lfloor \beta T_\beta^{i-1}(x) \rfloor$ for $i \geq 1$.

Rényi [27] showed that there is a unique probability measure μ_β on $[0, 1)$ that is equivalent to the Lebesgue measure and such that μ_β is invariant and ergodic with respect to T_β and has maximum entropy. The measure μ_β satisfies $(1 - \frac{1}{\beta})\lambda \leq \mu_\beta \leq \frac{\beta}{\beta-1}\lambda$.

Let $c(d)$ be the *cylinder set* corresponding to the block d , i.e. the set of all real numbers in the unit interval whose first $\|d\|$ digits coincide with d . A β -adic interval is a cylinder set $c(d)$ for some d .

Let W^∞ be the set of right-infinite words $\omega = \omega_1 \omega_2 \dots$ with digits $0 \leq \omega_i < \beta$ that appear as the β -expansions of real numbers in the unit interval. Let \mathcal{L}_n be the set of all finite subwords of length n of words $\omega \in W^\infty$ and let $W = \bigcup_{n \geq 1} \mathcal{L}_n$. We call the words in W *admissible*.

We have $\beta^n \leq |\mathcal{L}_n| \leq \frac{\beta}{\beta-1} \beta^n$ for the number of elements of \mathcal{L}_n .

For an infinite word $\omega = \omega_1 \omega_2 \dots \in W^\infty$ and a block $d = d_1 d_2 \dots d_k$ of digits $0 \leq d_i < \beta$ we denote by $N_d(\omega, n)$ the number of (possibly overlapping) occurrences of d within the first n letters of ω . If the word ω is finite, we write $N_d(\omega)$ for $N_d(\omega, \|\omega\|)$.

An infinite word $\omega \in W^\infty$ is called μ_β -normal if for all $d \in \mathcal{L}_k$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} N_d(\omega, n) = \mu_\beta(c(d)).$$

A real number $x \in [0, 1)$ is called *normal to base β* or β -normal, if the infinite word $\varepsilon_1 \varepsilon_2 \dots$ defined by its β -expansion (3.1) is μ_β -normal.

For fixed $\varepsilon > 0$ and positive integers k, n , a word $\omega \in \mathcal{L}_n$ is called (ε, k) -normal if for all $d \in \mathcal{L}_k$

$$\mu_\beta(c(d))(1 - \varepsilon)\|\omega\| < N_d(\omega) < \mu_\beta(c(d))(1 + \varepsilon)\|\omega\|.$$

The set of all (ε, k) -normal numbers in \mathcal{L}_n will be denoted by $E_n(\varepsilon, k)$ and its complement by $E_n^c(\varepsilon, k)$.

A *Pisot number* β is a real algebraic integer $\beta > 1$ such that all its conjugates have absolute value less than 1, and as usual we include all positive integers $b \geq 2$ in this definition. All Pisot numbers smaller than the golden mean were found by Dufresnoy and Pisot [17]. In particular, they showed that the smallest one is the positive root of $x^3 - x - 1$ (called the plastic number) which is approximately $1.32471 > \sqrt[3]{2}$.

3.2. Preliminaries.

Lemma 3.1 ([8, Lemma 3]). *Let $\beta > 1$ be Pisot. For every $\varepsilon > 0$ and positive integer k there exist $\eta = \eta(\varepsilon, k)$, $0 < \eta < 1$, $C = C(\varepsilon, k) > 0$ and $n_0 = n_0(\varepsilon, k)$ such that for the*

number of non- (ε, k) -normal words of length n

$$|E_n^c(\varepsilon, k)| < C |\mathcal{L}_n|^{1-\eta}$$

holds for all $n \geq n_0$.

In Section 4.2 we give explicit estimates for n_0 , C and η .

The following Lemma contains the underlying idea of our construction.

Lemma 3.2 ([8, Lemma 4]). *Let a_1, a_2, \dots be a sequence of finite words $a_n \in W$ such that $a = a_1 a_2 \dots \in W^\infty$ and $\|a_n\| \rightarrow \infty$ as $n \rightarrow \infty$. Suppose that for any $\varepsilon > 0$ and any positive integer k there exists an integer $n_0(\varepsilon, k)$ such that all a_n with $n \geq n_0(\varepsilon, k)$ are (ε, k) -normal. If*

$$(3.2) \quad n = o(\|a_1 a_2 \dots a_n\|) \quad \text{and} \quad \|a_{n+1}\| = o(\|a_1 a_2 \dots a_n\|),$$

then the infinite word $a = a_1 a_2 \dots$ is μ_β -normal.

Proof. Let $\varepsilon > 0$ and $d \in \mathcal{L}_k$. It suffices to show that, as $N \rightarrow \infty$,

$$\mu_\beta(c(d))(1 - \varepsilon)N < N_d(a, N) < \mu_\beta(c(d))(1 + \varepsilon)N.$$

We have $N_d(a, N) = N_d(a_1 a_2 \dots a_n, N)$, where n is such that $\|a_1 a_2 \dots a_{n-1}\| < N \leq \|a_1 \dots a_n\|$. Then, for N large enough,

$$\begin{aligned} N_d(a_1 \dots a_n, N) &\leq N_d(a_1 \dots a_{n_0(\varepsilon, k)}, N) + n(k-1) + N_d(a_{n_0+1}) + \dots + N_d(a_n) \\ &\leq \text{const}(\varepsilon, k) + n(k-1) + \sum_{i=n_0+1}^n \mu_\beta(c(d))(1 + \varepsilon)\|a_i\|. \end{aligned}$$

Dividing by N gives the desired result, assuming conditions (3.2). The calculation for the lower bound for $N_d(a, N)$ is similar. \square

Lemma 3.3. *Let $\beta > 1$ be Pisot. There exists $M \geq 0$ such that for all $n \geq 1$ and all $d \in L_n$ the Lebesgue measure of the cylinder set $c(d)$ satisfies*

$$(3.3) \quad \beta^{-(M+1)}\beta^{-n} \leq \lambda(c(d)) \leq \beta^{-n}.$$

Proof. This is Proposition 2.6 of [22]. \square

Following the argument in [22], one can take M to be the size of the largest block of consecutive zeros in the modified β -expansion of 1 (see Section 4.1). We give an explicit upper bound on M in Proposition 4.1.

We wish to control the lengths when changing the base. The following is an analogue to Lemma 3.3 in [6]; see also Lemma 2.1.

Lemma 3.4. *Let β be Pisot and M as above. For any interval I there is a β -adic subinterval I_β of I such that $\lambda(I_\beta) \geq \lambda(I)/2\beta^{M+4}$.*

Proof. We can assume $\lambda(I) > 0$. Let m be the smallest integer such that $\beta^{-m} < \lambda(I)$. Thus $\lambda(I)/\beta \leq \beta^{-m} < \lambda(I)$. If there exists an interval of order m in I , then let I_β be this β -adic interval and we have $\lambda(I_\beta) \geq \lambda(I)/\beta$.

Otherwise there must be a word $a \in \mathcal{L}_m$ such that $\pi(a) \in I$ but neither $\pi(a^-)$ nor $\pi(a^+)$ is in I , where a^- and a^+ are the lexicographically previous or next elements of a of the same length and where $\pi(a)$ is the real number in the unit interval whose β -expansion starts with a . Then by Lemma 3.3 we have that $\lambda(I) < 2\beta^{-m}$. Since $\beta^{-m} < \lambda(I)$ and the smallest Pisot number is bigger than $2^{1/3}$, we get that $2\beta^{-m-3} < \lambda(I)$. Thus there must be a β -adic interval I_β of order $m+3$ in I and we have

$$\lambda(I_\beta) \geq \frac{1}{\beta^{M+1+m+3}} = \frac{1}{2\beta^{M+4}} \cdot \frac{2}{\beta^m} > \frac{\lambda(I)}{2\beta^{M+4}}.$$

□

3.3. The Algorithm.

Notation. Let $(\beta_j)_{j \geq 1}$ be a sequence of Pisot numbers. Let t be a positive integer. A t -sequence is a sequence of intervals $\mathbf{I} = (I_1, \dots, I_t)$ such that for $1 \leq j \leq t$, I_j is β_j -adic, such that for $1 \leq j \leq t-1$, $I_{j+1} \subset I_j$, and such that $\lambda(I_{j+1}) \geq \lambda(I_j)/2\beta_{j+1}^{M_{\beta_{j+1}}+4}$. If we have two β -adic intervals $J \subset I$ then $u_\beta(J)$ means the block of digits that is added to the base β expansion of the numbers in I to obtain the β -expansion of numbers in J . The notation $u_j(\mathbf{J})$ for a t -sequence \mathbf{J} shall mean $u_{\beta_j}(J_j)$. We denoted the dependence on β_j of all appearing constants M , n_0 , C and of \mathcal{L}_n explicitly with an β_j .

Input. Given are values $\varepsilon_1 = 1$, $k_1 = 1$, $t_1 = 1$ and a sequence $(\beta_j)_{j \geq 1}$ of Pisot numbers β_j .

First step. Let \mathbf{I}_1 be a t_1 -sequence such that $\mathbf{I}_1 = (I_{1,t_1})$, with $I_{1,t_1} = [0, 1)$. Repeat the bases β_j according to conditions

$$(3.4) \quad \max_{1 \leq j \leq t_i} \beta_j \leq \beta_1 i,$$

$$(3.5) \quad \max_{1 \leq j \leq t_i} M_{\beta_j} \leq (M_{\beta_1} + 1)(1 + \log i),$$

$$(3.6) \quad \sum_{1 \leq j \leq t_i} (M_{\beta_j} + 4) \log \beta_j \leq (M_{\beta_1} + 4) \log \beta_1 (1 + \log i).$$

Step $i+1$ for $i \geq 1$. From step i , we have a t_i -sequence \mathbf{I}_i of nested intervals $I_{i,1} \supset \dots \supset I_{i,t_i}$ where each $I_{i,j}$ is β_j -adic.

Let

$$t_{i+1} = \lceil \log(i+1) \rceil, \quad \varepsilon_{i+1} = \frac{1}{t_{i+1}}, \quad k_{i+1} = t_{i+1},$$

$$\delta_{i+1} = \frac{1}{2} \frac{1}{2\beta_1^{M_{\beta_1}+4}} \frac{1}{t_i} \frac{1}{2^{t_i} \prod_{j \leq t_i} \beta_j^{M_{\beta_j}+4}} \frac{1}{2^{t_{i+1}} \prod_{j \leq t_{i+1}} \beta_j^{M_{\beta_j}+4}}.$$

Choose n_{i+1} to be the least integer such that

$$(3.7) \quad n_{i+1} \geq \max_{j \leq t_{i+1}} (n_{\beta_j}(\varepsilon_{i+1}, k_{i+1})),$$

and such that for all $1 \leq j \leq t_{i+1}$

$$(3.8) \quad \lambda(E_n^c(\varepsilon_{i+1}, k_{i+1})) < \delta_{i+1}.$$

Furthermore, let

$$v_i = \left\lceil \max_{j=1, \dots, t_i} \frac{\log \beta_j}{\log \beta_1} \right\rceil.$$

Then we perform the following steps.

- Take L to be a β_1 -adic interval of I_{i,t_i} of length $\lambda(L) \geq \lambda(I_{i,t_i}) 2^{-1} \beta_1^{-(M_{\beta_1}+4)}$.
- For each β_1 -adic sub-interval J_1 of L with $u_1(J_1) = v_i n_{i+1}$ find a t_{i+1} -sequence $\mathbf{J} = (J_1, \dots, J_{t_{i+1}})$.
- Choose the “leftmost” of the t_{i+1} -sequences \mathbf{J} such that $u_j(\mathbf{J})$ is $(\varepsilon_{i+1}, k_{i+1})$ -normal for $1 \leq j \leq t_i$.

Output. The unique real number X in the intersection of all $I_{i,j}$.

We need to show that the algorithm is well-defined and that the produced number is in fact β_j -normal for all $j \geq 1$.

Proposition 3.5. *This algorithm is well-defined.*

Proof. We have to show that in each step $i+1$ there exists at least one t_{i+1} -sequence \mathbf{J} . Let \mathcal{S} be the union of the intervals $J_{t_{i+1}}$ over the $|\mathcal{L}_{v_i n_{i+1}}^{\beta_1}|$ many t_{i+1} -sequences \mathbf{J} . By definition of the interval L we have that $\lambda(L) \geq \lambda(I_{i,t_i}) 2^{-1} \beta_1^{-(M_{\beta_1}+4)}$. Furthermore for each sequence we have that $\lambda(J_{t_{i+1}}) \geq 2^{-t_{i+1}} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(J_1)$. Since the sub-intervals $J_1 \subset L$ form a partition of L we have that $\lambda(\mathcal{S}) \geq 2^{-t_{i+1}} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(L)$. Combining these inequalities yields

$$\lambda(\mathcal{S}) \geq 2^{-t_i - t_{i+1} - 1} \prod_{j=1}^{t_i} \beta_j^{-(M_{\beta_j}+4)} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(I_{i,1}).$$

Now we calculate the measure of the set \mathcal{N} of non-suitable intervals and show that it is less than $\lambda(\mathcal{S})$. For the length of the added word we have $\|u_1(\mathbf{J})\| \geq v_i n_{i+1}$ and for each $2 \leq j \leq t_{i+1}$ we have $\|u_j(\mathbf{J})\| \geq n_{i+1}$. By the choice of n_{i+1} , the subsets of $I_{i,j}$, where $u_j(\mathbf{J})$ is not $(\varepsilon_{i+1}, k_{i+1})$ -normal, have Lebesgue measure less than $\delta_{i+1} \lambda(I_{i,j})$, and hence less than $\delta_{i+1} \lambda(I_{i,1})$. Since we consider t_i many bases, we obtain $\lambda(\mathcal{N}) < t_i \delta_{i+1} \lambda(I_{i,1})$.

Combining the estimates of \mathcal{N} and \mathcal{S} we obtain $\lambda(\mathcal{N}) < \lambda(\mathcal{S})$. Since $\mathcal{N} \subset \mathcal{S}$ there must be a t_{i+1} -sequence \mathbf{J} such that $u_j(\mathbf{J})$ is $(\varepsilon_{i+1}, k_{i+1})$ -normal for each $1 \leq j \leq t_i$. \square

Theorem 3.6. *Let $(\beta_j)_{j \geq 1}$ be a sequence of Pisot numbers. Then the real number X generated by this algorithm is β_j -normal for each $j \geq 1$.*

Proof. We need to verify the growth and normality assumptions of Lemma 3.2 on the words that correspond to the digits added in each considered base in each step of the algorithm.

To find bounds for the number of added digits in step $i+1$ in base β_j , for $j \leq t_i$, consider the chain of intervals

$$I_{i,j} \supset \dots \supset I_{i,t_i} \supset L \supset J_1 \supset \dots \supset J_j$$

which is considered in step $i+1$. We find a lower bound on the Lebesgue measure of J_j in the form of

$$\lambda(J_j) \geq \frac{1}{2^{t_i}} \frac{1}{\beta_1^{M_1+1}} \frac{1}{\beta_1^{v_i n_{i+1}}} \prod_{l=1}^{t_i} \frac{1}{\beta_l^{M_{\beta_l}+4}} \cdot \lambda(I_{i,j}).$$

Thus, Lemma 3.3 implies for the number $\|u_j^{(i+1)}(\mathbf{J})\|$ of digits added in base β_j , $j \leq t_i$, in step $i+1$ of the algorithm, that

$$\frac{\log \left(\frac{1}{F_{i+1}} \frac{1}{\beta_j^{M_{\beta_j}+1}} \right)}{\log \beta_j} \leq \|u_j^{(i+1)}(\mathbf{J})\| \leq \frac{\log \frac{1}{F_{i+1}}}{\log \beta_j}$$

where $F_{i+1} = 2^{-t_i} \beta_1^{-(M_1+1)} \beta_1^{-v_i n_{i+1}} \prod_{l=1}^{t_i} \beta_l^{-M_{\beta_l}-4}$.

Hence $\|u_j^{(i+1)}(\mathbf{J})\| \sim \log 1/F_{i+1}$ with implied constants only depending on β_j . We thus need to show that

$$\log 1/F_{i+1} = t_i \log 2 + (v_i n_{i+1} + M_1 + 1) \log \beta_1 + \sum_{l=1}^{t_i} (M_l + 4) \log \beta_l$$

satisfies assumptions (3.2) of Lemma 3.2.

We now look at the growth of n_{i+1} . In light of Proposition 4.3, condition (3.7) requires

$$(3.9) \quad n_{i+1} \geq M_{\beta_j} + k_{i+1}$$

for all $1 \leq j \leq t_{i+1}$. We have $\varepsilon_{i+1} = 1/t_{i+1} \rightarrow 0$ and $k_{i+1} = t_{i+1} \rightarrow \infty$ as $t_{i+1} \rightarrow \infty$. Thus also n_{i+1} tends to infinity at least logarithmically in i .

Since $\lambda \leq \frac{\beta}{\beta-1} \mu_\beta$, $\beta^k \leq |\mathcal{L}_k| \leq \frac{\beta}{\beta-1} \beta^k$, and because of Proposition 4.3, condition (3.8) on n_{i+1} is satisfied, if for all $j \leq t_i$,

$$4 \left(\frac{\beta_j}{\beta_j - 1} \right)^2 \beta_j^k \beta_j^{n_{i+1} \eta(\varepsilon_{i+1}, k_{i+1})} < \delta_{i+1}.$$

With η from equation (4.3), this translates into the requirement that for every $j \leq t_i$,

$$(3.10) \quad n_{i+1} \geq \frac{(M_{\beta_j} + 1) \log \beta_j + \log \frac{\beta_j}{\beta_j - 1}}{\varepsilon_{i+1} \min(\frac{\varepsilon_{i+1} \beta_j^{k_{i+1}}}{16}, \frac{3}{4})} \left(\log \left(4 \left(\frac{\beta_j}{\beta_j - 1} \right)^2 \beta_j^{k_{i+1}} \right) + \log \frac{1}{\delta_{i+1}} \right),$$

where

$$\begin{aligned} \log \frac{1}{\delta_{i+1}} = & 2 \log 2 + \log t_i + (t_i + t_{i+1}) \log 2 + (M_{\beta_1} + 4) \log \beta_1 \\ & + 2 \sum_{1 \leq j \leq t_i} (M_{\beta_j} + 4) \log \beta_j + \sum_{t_i < j \leq t_{i+1}} (M_{\beta_j} + 4) \log \beta_j \end{aligned}$$

(where the last sum is empty if $t_i = t_{i+1}$).

Properties (3.4) and (3.6) on the sequence $(\beta_j)_{j \geq 1}$ imply
(3.11)

$$\max_{1 \leq j \leq t_i} \left((M_{\beta_j} + 1) \log \beta_j + \log \frac{\beta_j}{\beta_j - 1} \right) \leq \left((M_{\beta_1} + 4) \log \beta_1 + \log \frac{\beta_1}{\sqrt[3]{2} - 1} + 1 \right) (1 + \log i).$$

Conditions (3.4) - (3.6) can be achieved by suitably repeating the bases β_j . All conditions are satisfied in step 1, and the process of repeating the bases is possible computably.

Properties (3.4) - (3.6) and (3.11), together with $t_{i+1} = k_{i+1} = 1/\varepsilon_{i+1} \sim \log i$, imply that for i large enough

$$n_{i+1} \geq O \left(\frac{\log i}{1/\log i} (\log i + (\log i)^2 + \log \log i + \log i + \log i) \right) = O((\log i)^4)$$

where the implied constant only depends on β_1 . Hence n_{i+1} grows at least as $O(\log i)$ and at most as $O((\log i)^4)$, where the implied constants depend only on β_1 . Thus $\log 1/F_{i+1}$ and hence also $\|u_j^{(i+1)}(\mathbf{J})\|$ grows at least as $O(\log i)$ and at most as $O((\log i)^4)$, where again the implied constants only depend on β_1 . Thus $\|u_j^{(i+1)}(\mathbf{J})\|$ satisfies conditions (3.2) of Proposition 3.2. Hence the number X produced by this algorithm is β_j -normal for every $j \geq 1$. \square

Remark. The choices of how t_i , ε_i and k_i change with the step i of the algorithm and the conditions on the sequence of bases $(\beta_j)_{j \geq 1}$ are rather arbitrary. There is a lot of freedom to optimize for other quantities, such as done in Becher, Heiber, Slaman [6] where computational speed is optimized. This is not taken into account here.

Remark. Following these lines, an extension of Becher, Heiber, Slaman's algorithm to a countable set of real bases that are β -numbers is possible, provided these bases are bounded away from 1 and such there is a uniform bound on the length of the periodic part in their orbit of 1.

A β -number is a real number β such that the orbit of 1 under T_β is finite. Pisot numbers are β -numbers. It is not known under which conditions Salem numbers are or are not β -numbers (a *Salem number* is a real algebraic integer $\beta > 1$ such that all its conjugates have absolute values at most equal to one, with equality in at least one case). Salem numbers of degree 4 are β -numbers, but there is computational and heuristic evidence that higher degree Salem numbers exist that are no β -numbers, see for example [12].

Note that β -numbers satisfy the specification property - one can always use a block of zeros to make the concatenation of two admissible blocks admissible. This is because admissible words can be characterized as precisely the subwords of the lexicographic largest

word in the β -shift. Since the orbit of 1 is finite, this word will be eventually periodic and hence the lengths of subwords consisting of only zeros is bounded. Thus Lemma 3 in [8] on the number of (ε, k) -normal admissible words is valid and can be used as an existence criterion for a t_i sequence \mathbf{J} in each step of the algorithm.

Note also that β -numbers also satisfy Proposition 2.6 of [22] needed to control the decay of the length of subintervals. However, we are looking for a lower bound for the measure of cylinder intervals of the form (3.3) that is uniform for all bases β under consideration. This can be achieved by requiring that there is a uniform bound on the length of the period of the orbit of 1 under T_β for each β under consideration.

When adapting the proof of Lemma 3.4 to β -numbers, we moreover need to require that the set of β -numbers under consideration is bounded away from 1, as above with the plastic number.

4. EXPLICIT ESTIMATES FOR β -EXPANSIONS

In this section we make explicit the constants in Lemma 3.1 using large deviation estimates for certain dependent random variables. This requires us to provide an upper bound for the length of the largest block of zeros appearing in the modified β -expansion of 1 for a Pisot number β .

4.1. Number of zeros in the expansion of 1. Let β be a Pisot number and denote by $d_\beta(1) = 0.\varepsilon_1\varepsilon_2\dots$ the β -expansion of 1, i.e. $\varepsilon_1 = \lfloor \beta \rfloor$ and $\varepsilon_i = \lfloor \beta T_\beta^{i-1}(1) \rfloor$ for $i \geq 1$. Let $d_\beta^*(1)$ be the modified β -expansion of 1, i.e. $d_\beta^*(1) = d_\beta(1)$ if the sequence $\varepsilon_1\varepsilon_2\dots$ does not end with infinitely many zeros, and $d_\beta^*(1) = 0.(\varepsilon_1\varepsilon_2\dots\varepsilon_{n-1}(\varepsilon_n - 1))^\omega$ when $d_\beta(1)$ ends in infinitely many zeros and ε_n is the last non-zero digit. It is known that $d_\beta^*(1)$ is purely periodic or eventually periodic if β is Pisot. We reprove this fact here and give an explicit upper bound for the preperiod length v and period length p and take $v + p$ as a trivial upper bound for the size of the largest block of zeros in $d_\beta^*(1)$. Note that $d_\beta^*(1)$ is (eventually) periodic if the orbit of 1 under T_β is finite, and that the number of distinct elements in this orbit is precisely $v + p$.

Proposition 4.1. *Let β be a Pisot number of degree d with r real conjugates $\beta = \beta_1, \beta_2, \dots, \beta_r$ and $2s$ complex conjugates $\beta_{r+1}, \dots, \beta_d$. Then the orbit of 1 under the map T_β , i.e. the set*

$$\{T_\beta^k(1) \mid k \geq 0\},$$

is finite and its number of elements is bounded by

$$(4.1) \quad M = d! \det(B)^{-1} 2^{r+s-1} \pi^s C^{r+2s-1} + d$$

where

$$(4.2) \quad B = \begin{pmatrix} 1 & \beta & \dots & \beta^{d-1} \\ 1 & \beta_2 & \dots & \beta_2^{d-1} \\ \vdots & \vdots & & \vdots \\ 1 & \beta_d & \dots & \beta_d^{d-1} \end{pmatrix}$$

and where

$$C = 1 + \frac{\lfloor \beta \rfloor}{1 - \eta}$$

with $\eta = \max_{2 \leq j \leq d} |\beta_j| < 1$.

Proof. For $k \geq 0$, $T_\beta^k(1)$ is an element of $\mathbb{Z}[\beta]$, hence there is a unique representation $T_\beta^k(1) = p_0^{(k)} + p_1^{(k)}\beta + \dots + p_{d-1}^{(k)}\beta^{d-1}$ with $p_i^{(k)} \in \mathbb{Z}$. Denote by σ_j , $1 \leq j \leq d$, the j -th conjugation, ordered such that the first r are real, and $\sigma_{r+i} = \bar{\sigma}_{r+s+i}$ for $1 \leq i \leq s$. We have

$$T_\beta^k(1) = \beta^k \left(1 - \sum_{l=1}^k \varepsilon_l \beta^{-l} \right)$$

hence for $2 \leq j \leq d$

$$|\sigma_j(T_\beta^k(1))| \leq 1 + \frac{\lfloor \beta \rfloor}{1 - \eta}$$

where $\eta = \max_{2 \leq j \leq d} |\beta_j| < 1$.

Note that

$$B \begin{pmatrix} p_0^{(k)} \\ p_1^{(k)} \\ \vdots \\ p_{d-1}^{(k)} \end{pmatrix} = \begin{pmatrix} T_\beta^k(1) \\ \sigma_2(T_\beta^k(1)) \\ \vdots \\ \sigma_d(T_\beta^k(1)) \end{pmatrix}$$

where B is as in (4.2) and has determinant $\det B = \prod_{1 \leq i < j \leq d} (\beta_j - \beta_i) \neq 0$. Now, since the vector of $T_\beta^k(1)$ and its conjugates can be canonically embedded in a compact convex set in \mathbb{R}^{r+2s} of volume $2^{r+s-1} \pi^s C^{r+2s-1}$, we can count the \mathbb{Z}^d -lattice points in a compact convex set in \mathbb{R}^d of volume $\det(B)^{-1} 2^{r+s-1} \pi^s C^{r+2s-1}$. By loosing a factor of 2, we can make this set additionally centrally symmetric if we allow $T_\beta^k(1)$ (formally) to take on values in the interval $[-1, 1]$. Then we can use a result by Blichfeldt [10] and bound the number of \mathbb{Z}^d -lattice points in $B^{-1}Y$ by

$$|B^{-1}Y \cap \mathbb{Z}^d| \leq d! \det(B)^{-1} 2^{r+s-1} \pi^s C^{r+2s-1} + d$$

with $C = 1 + \frac{\lfloor \beta \rfloor}{1 - \eta}$ and hence obtain an upper bound for the number of distinct points in the orbit of 1 under T_β which is also a trivial upper bound for the maximum number of consecutive zeros in the modified β -expansion of 1 as explained above. \square

4.2. Number of not (ε, k) -normal numbers. Let β be a Pisot number and let \mathcal{L}_n be the set of all admissible words of length n . Fix $\varepsilon > 0$ and a positive integer k . We wish to find explicit estimates for the number of non- (ε, k) -normal words of length n for fixed $\varepsilon > 0$ and k such as given in Lemma 3.1 (Lemma 3 in [8]). The method in [8] uses methods of ergodic theory and the authors are not aware of a method to make the implied constants explicit. Therefore we use a probabilistic approach by viewing the digits to base β as random variables and using a variant of Hoeffding's inequality for dependent random variables to bound the tail distribution of their sum. This approach automatically gives

all involved constants explicitly. We use the following Lemma due to Siegel (Theorem 5 in [34]).

Lemma 4.2. *Let $X = X_1 + X_2 + \dots + X_l$ be the sum of l possibly dependent random variables. Suppose that X_i , for $i = 1, 2, \dots, l$, is the sum of n_i mutually independent random variables having values in the interval $[0, 1]$. Let $\mathbb{E}[X_i] = n_i p_i$. Then for $a \geq 0$*

$$\mathbb{P}(X - \mathbb{E}[X] \geq a) < \exp\left(-\frac{a^2}{8(\sum_i \sqrt{p_i(1-p_i)n_i})^2}\right) + \exp\left(-\frac{3a}{4\sum_i(1-p_i)^2}\right).$$

Proposition 4.3. *Let β be a Pisot number. The μ_β -measure of the set of not (ε, k) -normal words of length n satisfies*

$$\mu_\beta(E_n^c(\varepsilon, k)) \leq 4|\mathcal{L}_k||\mathcal{L}_n|^{-\eta}$$

for $n \geq M + k$ with $\eta > 0$ as in equation (4.3) and M as in equation (4.1).

Proof. Let $d \in \mathcal{L}_k$ and for $n \geq M + k$, let $X_1, \dots, X_{M+1} : \mathcal{L}_n \rightarrow \mathbb{R}$ be random variables where $X_i(\omega)$ denotes the number of occurrences of the word d in $\omega = \omega_1 \dots \omega_n$ at positions

$$\omega_{i+j(M+1)}\omega_{i+j(M+1)+1} \dots \omega_{i+j(M+1)+k-1}$$

for $0 \leq j \leq \lfloor \frac{n-k}{M+1} \rfloor$. The X_i are dependent, but each is a sum of $n_i = \lfloor \frac{n-k}{M+1} \rfloor + 1$ independent identically distributed random variables $Y_j^{(i)}$ that take value one if and only if the word d appears in ω starting at digit $\omega_{i+j(M+1)}$ and zero otherwise. We have $\mathbb{E}[X] = n\mu_\beta(c(d))$ and $\mathbb{E}[X_i] = n_i\mu_\beta(c(d))$. Denote by $\bar{E}_n(\varepsilon, k)$ the set of words of length n for which there is a subword d of length k that appears more often than $n(\mu_\beta(c(d)) + \varepsilon)$ times and let $\bar{E}_n(\varepsilon, d)$ be the set of words of length n for which the subword d appears more often than $n(\mu_\beta(c(d)) + \varepsilon)$ times. We apply Lemma 4.2 with $l = M + 1$, n_i as above, $p_i = \mu_\beta(c(d))$ and $a = n\varepsilon$ and obtain

$$\begin{aligned} \mu_\beta(X > n(\mu_\beta(c(d)) + \varepsilon)) &= \mu_\beta(\bar{E}_n(\varepsilon, d)) \\ &< \exp\left(-\frac{(n\varepsilon)^2}{8\mu_\beta(c(d))(1-\mu_\beta(c(d)))(M+1)^2(\lfloor \frac{n-k}{M+1} \rfloor + 1)}\right) \\ &\quad + \exp\left(-\frac{3n\varepsilon}{4(M+1)(1-\mu_\beta(c(d)))^2}\right). \end{aligned}$$

Using $c\beta^{-k} \leq \mu_\beta(c(d)) \leq \beta^{-k}$ and $n \geq M + 1$, this is

$$< \exp\left(-\frac{\varepsilon^2 n}{16(M+1)\beta^{-k}}\right) + \exp\left(-\frac{3\varepsilon n}{4(M+1)}\right) < 2 \exp\left(-\frac{\varepsilon n}{M+1} \min\left(\frac{\varepsilon}{16\beta^{-k}}, \frac{3}{4}\right)\right).$$

Finally, since $\mu_\beta(\bar{E}_n(\varepsilon, k)) \leq \sum_{d \in \mathcal{L}_k} \mu_\beta(\bar{E}_n(\varepsilon, d))$ and using that $\beta^n \leq |\mathcal{L}_n| \leq \frac{\beta}{\beta-1}\beta^n$ we obtain

$$\begin{aligned} \mu_\beta(\bar{E}_n(\varepsilon, k)) &\leq |\mathcal{L}_k| 2 \exp\left(-\frac{\varepsilon n}{M+1} \min\left(\frac{\varepsilon}{16\beta^{-k}}, \frac{3}{4}\right)\right) \\ &\leq 2|\mathcal{L}_k||\mathcal{L}_n|^{-\eta} \end{aligned}$$

with

$$(4.3) \quad \eta = \frac{\varepsilon \min(\frac{\varepsilon}{16\beta^{-k}}, \frac{3}{4})}{\log(\frac{\beta}{\beta-1}) + (M+1) \log \beta} > 0.$$

Using the same argument with $Y = n - X$ gives a symmetrical upper bound for the number of words ω of length n in which the word d appears less than $n\mu_\beta(c(d)) - \varepsilon n$ times. Thus we obtain an upper bound for the number of not (ε, k) -normal words of length n of the form

$$4|\mathcal{L}_k| |\mathcal{L}_n|^{-\eta}$$

for $n \geq M + k$ with η as in (4.3). □

Corollary 4.4. *The number of not (ε, k) -normal words of length n satisfies*

$$|E_n^c(\varepsilon, k)| \leq C |\mathcal{L}_n|^{1-\eta}$$

for $n \geq M + k$ with $\eta > 0$ as in equation (4.3), M as in equation (4.1), and where $C = 4|\mathcal{L}_k| \beta^{M+1} \frac{\beta}{\beta-1}$.

Proof. Since the Parry measure μ_β satisfies

$$\left(1 - \frac{1}{\beta}\right) \lambda \leq \mu_\beta \leq \frac{\beta}{\beta-1} \lambda$$

with respect to the Lebesgue measure λ , and due to the bounds on the Lebesgue measure of β -adic cylinder intervals from Lemma 3.3, the bound from Proposition 4.3 on the μ_β measure of the set of non- (ε, k) -normal words of length n implies for the number of such words

$$(4.4) \quad |E_n^c(\varepsilon, k)| \leq C |\mathcal{L}_n|^{1-\eta},$$

where $C = C(\beta, k) = 4|\mathcal{L}_k| \beta^M \frac{\beta}{\beta-1}$ and $\eta = \eta(\beta, \varepsilon, k)$ as given in equation (4.3) and where we used that $\beta^n \leq |\mathcal{L}_n| \leq \frac{\beta}{\beta-1} \beta^n$. □

Acknowledgements. For the realization of the present paper the first author received support from the Conseil Régional de Lorraine. Parts of this research work were done when the first author was visiting the Department of Mathematics of Graz University of Technology. The author thanks the institution for their hospitality. The second author was supported by the Austrian Science Fund (FWF): I 1751-N26; W1230, Doctoral Program “Discrete Mathematics”; and SFB F 5510-N26. He would like to thank Karma Dajani and Bing Li for some interesting discussions on β -expansions.

REFERENCES

- [1] N. Alvarez and V. Becher, *M. Levin's construction of absolutely normal numbers with very low discrepancy*, arXiv:1510.02004.
- [2] D. H. Bailey and J. Borwein, *Pi Day is upon us again and we still do not know if pi is normal*, Amer. Math. Monthly **121** (2014), no. 3, 191–206, DOI 10.4169/amer.math.monthly.121.03.191.
- [3] V. Becher, Y. Bugeaud, and T. A. Slaman, *On Simply Normal Numbers to Different Bases*, arXiv:1311.0332 (2013).
- [4] V. Becher and S. Figueira, *An example of a computable absolutely normal number*, Theoretical Computer Science **270** (2002), 126–138.
- [5] V. Becher, S. Figueira, and R. Picchi, *Turing's unpublished algorithm for normal numbers*, Theoret. Comput. Sci. **377** (2007), no. 1-3, 126–138, DOI 10.1016/j.tcs.2007.02.022.
- [6] V. Becher, P. A. Heiber, and T. A. Slaman, *A polynomial-time algorithm for computing absolutely normal numbers*, Inform. and Comput. **232** (2013), 1–9. MR3132518
- [7] V. Becher and T. A. Slaman, *On the normality of numbers to different bases*, J. Lond. Math. Soc. (2) **90** (2014), no. 2, 472–494, DOI 10.1112/jlms/jdu035.
- [8] A. Bertrand-Mathis and B. Volkmann, *On (ε, k) -normal words in connecting dynamical systems*, Monatsh. Math. **107** (1989), no. 4, 267–279. MR1012458 (90m:11115)
- [9] A. S. Besicovitch, *The asymptotic distribution of the numerals in the decimal representation of the squares of the natural numbers*, Math. Zeit. **39** (1935), 146–156.
- [10] H. F. Blichfeldt, *Notes on Geometry of Numbers*, Bull. Amer. Math. Soc. **27** (1921), no. 4, 150–153.
- [11] É. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rendiconti del Circolo Matematico di Palermo **27** (1909), no. 1, 247–271, DOI 10.1007/BF03019651 (French).
- [12] D. W. Boyd, *On the beta expansion for Salem numbers of degree 6*, Math. Comp. **65** (1996), no. 214, 861–875, S29–S31, DOI 10.1090/S0025-5718-96-00700-4.
- [13] Y. Bugeaud, *Distribution Modulo One and Diophantine Approximation*, Cambridge Tracts in Mathematics, Cambridge University Press, 2012.
- [14] G. J. Chaitin, *A theory of program size formally identical to information theory*, J. Assoc. Comput. Mach. **22** (1975), 329–340.
- [15] D. G. Champernowne, *The Construction of Decimals Normal in the Scale of Ten*, J. London Math. Soc. **S1-8**, no. 4, 254, DOI 10.1112/jlms/s1-8.4.254.
- [16] H. Davenport and P. Erdős, *Note on normal decimals*, Canadian J. Math. **4** (1952), 58–63.
- [17] J. Dufresnoy and Ch. Pisot, *Etude de certaines fonctions méromorphes bornées sur le cercle unité. Application à un ensemble fermé d'entiers algébriques*, Ann. Sci. Ecole Norm. Sup. (3) **72** (1955), 69–92. MR0072902 (17,349d)
- [18] K. Fukuyama, *Metric discrepancy results for alternating geometric progressions*, Monatsh. Math. **171** (2013), no. 1, 33–63, DOI 10.1007/s00605-012-0419-4.
- [19] S. Gaal and L. Gál, *The discrepancy of the sequence $\{(2^n x)\}$* , Nederl. Akad. Wetensch. Proc. Ser. A **67** = Indag. Math. **26** (1964), 129–143.
- [20] M. B. Levin, *Absolutely normal numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. **1** (1979), 31–37, 87.
- [21] ———, *On the discrepancy estimate of normal numbers*, Acta Arith. **88** (1999), no. 2, 99–111.
- [22] B. Li and J. Wu, *Beta-expansion and continued fraction expansion*, J. Math. Anal. Appl. **339** (2008), no. 2, 1322–1331. MR2377089 (2008m:11148)
- [23] M. G. Madritsch, J. M. Thuswaldner, and R. F. Tichy, *Normality of numbers generated by the values of entire functions*, J. Number Theory **128** (2008), no. 5, 1127–1145, DOI 10.1016/j.jnt.2007.04.005.
- [24] Y. Nakai and I. Shiokawa, *A class of normal numbers*, Japan. J. Math. (N.S.) **16** (1990), no. 1, 17–29.
- [25] W. Parry, *On the β -expansions of real numbers*, Acta Math. Acad. Sci. Hungar. **11** (1960), 401–416.
- [26] W. Philipp, *Limit theorems for lacunary series and uniform distribution mod 1*, Acta Arith. **26** (1974/75), no. 3, 241–251.

- [27] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar **8** (1957), 477–493.
- [28] A.-M. Scheerer, *Computable Absolutely Normal Numbers and Discrepancies*, arXiv:1511.03582 (2015).
- [29] ———, *Normality in Pisot Numeration Systems*, arXiv:1503.08047 (2015).
- [30] J. Schiffer, *Discrepancy of normal numbers*, Acta Arith. **47** (1986), no. 2, 175–186.
- [31] W. M. Schmidt, *On normal numbers*, Pacific J. Math. **10** (1960), 661–672.
- [32] ———, *Über die Normalität von Zahlen zu verschiedenen Basen*, Acta Arith. **7** (1961/1962), 299–309.
- [33] ———, *Irregularities of distribution. VII*, Acta Arith. **21** (1972), 45–50.
- [34] A. Siegel, *Toward a usable theory of Chernoff Bounds for heterogeneous and partially dependent random variables* (1992).
- [35] W. Sierpinski, *Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre*, Bulletin de la Société Mathématique de France **45** (1917), 127–132.
- [36] A. Turing, *A Note on Normal Numbers*, Collected Works of A. M. Turing, Pure Mathematics, edited by J. L. Britton, 1992, pp. 117–119.
- [37] D. D. Wall, *Normal Numbers* (1950). Thesis (Ph.D.)—University of California, Berkeley.

(M. G. Madritsch) 1. UNIVERSITÉ DE LORRAINE, INSTITUT ELIE CARTAN DE LORRAINE, UMR 7502, VANDOEUVRE-LÈS-NANCY, F-54506, FRANCE;

2. CNRS, INSTITUT ELIE CARTAN DE LORRAINE, UMR 7502, VANDOEUVRE-LÈS-NANCY, F-54506, FRANCE

E-mail address: `manfred.madritsch@univ-lorraine.fr`

(A.-M. Scheerer) DEPARTMENT FOR ANALYSIS AND COMPUTATIONAL NUMBER THEORY, GRAZ UNIVERSITY OF TECHNOLOGY, A-8010 GRAZ, AUSTRIA

E-mail address: `scheerer@math.tugraz.at`

(R. F. Tichy) DEPARTMENT FOR ANALYSIS AND COMPUTATIONAL NUMBER THEORY, GRAZ UNIVERSITY OF TECHNOLOGY, A-8010 GRAZ, AUSTRIA

E-mail address: `tichy@tugraz.at`